

# André Mourão

R&D with experience in Information Retrieval, Computer Vision and Machine Learning

[Personal Site](#)

[GitHub](#)

[LinkedIn](#)

[andre.b.mourao@gmail.com](mailto:andre.b.mourao@gmail.com)

+351917915979

## EXPERIENCE

- 2020 – present | **R&D Engineer at the Portuguese Web Archive** [arquivo.pt](http://arquivo.pt)  
- **Responsible** for page, image search and user-facing APIs  
- **Increased Solr** image metadata search index **90-fold**, from 20 million to **1.8 billion docs**
- 2019 – present | **Freelancer:**  
- **moetsi:** designed a Distributed Video Processing system for Computer Vision at scale
- 2018 | **Post-doc researcher at the COGNITUS Horizon 2020 project** [now.novasearch.org](http://now.novasearch.org)  
- Created the event plotting module, enabling journalists to find the best photos and videos for news articles. It leverages on social media trends using **NLP** and **unsupervised ML**
- 2015 | **Visiting Researcher at the University of Texas, Austin**  
Extraction and analysis of smartwatch activity and health data.

## TECHNOLOGIES

- Technologies | **IR/Search** (Java, Solr, Elasticsearch, Lucene, Hadoop)  
**Computer Vision** (C++, OpenCV, LibAV, LibSVM, FLANN)  
**Data Science and NLP** (Python, NumPy, SciPy, sklearn, SpaCy Django)
- Portfolio | **NovaMedSearch** [medical.novasearch.org](http://medical.novasearch.org)  
- Search engine using both images and text queries to find cases in the biomedical literature.  
- Creator of the image search, medical text annotation and clinical trial search systems
- Revisionista.pt** [revisionista.pt](http://revisionista.pt)  
- Find out how news articles are edited after publication across Portuguese newsrooms

## EDUCATION

- 2012 – 2018 | **Ph.D. in Computer Science** *NOVA SST* Supervisor: Prof. João Magalhães  
- Developed an architecture to improve the partitioning of very large scale (over 1 billion) image search for distributed systems using Matrix Factorization and Clustering.  
- Benchmarked on Microsoft Azure cloud, with a difference in load of <10% between nodes.
- 2010 – 2012 | **M.Sc. degree in Computer Science** *NOVA School of Science and Technology*  
- Created a SVM model and image features for real-time facial expression recognition  
- The model was successfully applied on a competitive game based solely on player's faces.
- 2007 – 2010 | **B.Sc. degree in Computer Science** *NOVA School of Science and Technology*

## AWARDS AND DISTINCTIONS

- 2019 | **Arquivo.pt award** – 2<sup>nd</sup> place for the **Revisionista.pt** project
- 2015 – 2017 | **Microsoft Azure Research Award** – Awarded **\$40.000** in Microsoft Azure credit to apply on improving index partitioning for distributed search of biomedical multimodal data.
- 2013 | **Fraunhofer Portugal Challenge** Awarding Research of Practical Utility – 2<sup>nd</sup> place: **M.Sc.**